

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/580,458A
Source: IFWO
Date Processed by STIC: 8/22/06

ENTERED



IFWO

RAW SEQUENCE LISTING

DATE: 08/22/2006

PATENT APPLICATION: US/10/580,458A

TIME: 10:54:29

Input Set : A:\65691-445CorrSeqList.txt

Output Set: N:\CRF4\08222006\J580458A.raw

```

3 <110> APPLICANT: Assistance Publique - Hopitaux de Paris (AH-HP)
4   Institut National de la Sante et de la Recherche Medicale
5   (INSERM)
6   Institut Gustave Roussy (IGR)
7   Universite de Versailles - Saint-Quentin-en-Yvelines
8   Universite Paris-Sud
9   VAINCHENKER, William
10  UGO, Valerie
11  JAMES, Chloe
12  LE COUEDIC, Jean-Pierre
13  CASADEVALL, Nicole
15 <120> TITLE OF INVENTION: Identification of a JAK2 mutation involved in Vaquez
16   Polyglobulia
18 <130> FILE REFERENCE: D 22707
C--> 20 <140> CURRENT APPLICATION NUMBER: US/10/580,458A
C--> 21 <141> CURRENT FILING DATE: 2006-05-24
23 <160> NUMBER OF SEQ ID NOS: 31
25 <170> SOFTWARE: PatentIn version 3.3
27 <210> SEQ ID NO: 1
28 <211> LENGTH: 1132
29 <212> TYPE: PRT
30 <213> ORGANISM: homo sapiens
33 <220> FEATURE:
34 <223> OTHER INFORMATION: variant JAK2 V617F
36 <400> SEQUENCE: 1
38 Met Gly Met Ala Cys Leu Thr Met Thr Glu Met Glu Gly Thr Ser Thr
39 1          5          10          15
41 Ser Ser Ile Tyr Gln Asn Gly Asp Ile Ser Gly Asn Ala Asn Ser Met
42          20          25          30
44 Lys Gln Ile Asp Pro Val Leu Gln Val Tyr Leu Tyr His Ser Leu Gly
45          35          40          45
47 Lys Ser Glu Ala Asp Tyr Leu Thr Phe Pro Ser Gly Glu Tyr Val Ala
48          50          55          60
50 Glu Glu Ile Cys Ile Ala Ala Ser Lys Ala Cys Gly Ile Thr Pro Val
51 65          70          75          80
53 Tyr His Asn Met Phe Ala Leu Met Ser Glu Thr Glu Arg Ile Trp Tyr
54          85          90          95
56 Pro Pro Asn His Val Phe His Ile Asp Glu Ser Thr Arg His Asn Val
57          100         105         110
59 Leu Tyr Arg Ile Arg Phe Tyr Phe Pro Arg Trp Tyr Cys Ser Gly Ser
60          115         120         125
62 Asn Arg Ala Tyr Arg His Gly Ile Ser Arg Gly Ala Glu Ala Pro Leu
63          130         135         140

```

RAW SEQUENCE LISTING

DATE: 08/22/2006

PATENT APPLICATION: US/10/580,458A

TIME: 10:54:29

Input Set : A:\65691-445CorrSeqList.txt

Output Set: N:\CRF4\08222006\J580458A.raw

```

65 Leu Asp Asp Phe Val Met Ser Tyr Leu Phe Ala Gln Trp Arg His Asp
66 145                150                155                160
68 Phe Val His Gly Trp Ile Lys Val Pro Val Thr His Glu Thr Gln Glu
69                165                170                175
71 Glu Cys Leu Gly Met Ala Val Leu Asp Met Met Arg Ile Ala Lys Glu
72                180                185                190
74 Asn Asp Gln Thr Pro Leu Ala Ile Tyr Asn Ser Ile Ser Tyr Lys Thr
75                195                200                205
77 Phe Leu Pro Lys Cys Ile Arg Ala Lys Ile Gln Asp Tyr His Ile Leu
78                210                215                220
80 Thr Arg Lys Arg Ile Arg Tyr Arg Phe Arg Arg Phe Ile Gln Gln Phe
81 225                230                235                240
83 Ser Gln Cys Lys Ala Thr Ala Arg Asn Leu Lys Leu Lys Tyr Leu Ile
84                245                250                255
86 Asn Leu Glu Thr Leu Gln Ser Ala Phe Tyr Thr Glu Lys Phe Glu Val
87                260                265                270
89 Lys Glu Pro Gly Ser Gly Pro Ser Gly Glu Glu Ile Phe Ala Thr Ile
90                275                280                285
92 Ile Ile Thr Gly Asn Gly Gly Ile Gln Trp Ser Arg Gly Lys His Lys
93                290                295                300
95 Glu Ser Glu Thr Leu Thr Glu Gln Asp Leu Gln Leu Tyr Cys Asp Phe
96 305                310                315                320
98 Pro Asn Ile Ile Asp Val Ser Ile Lys Gln Ala Asn Gln Glu Gly Ser
99                325                330                335
101 Asn Glu Ser Arg Val Val Thr Ile His Lys Gln Asp Gly Lys Asn Leu
102                340                345                350
104 Glu Ile Glu Leu Ser Ser Leu Arg Glu Ala Leu Ser Phe Val Ser Leu
105                355                360                365
107 Ile Asp Gly Tyr Tyr Arg Leu Thr Ala Asp Ala His His Tyr Leu Cys
108                370                375                380
110 Lys Glu Val Ala Pro Pro Ala Val Leu Glu Asn Ile Gln Ser Asn Cys
111 385                390                395                400
113 His Gly Pro Ile Ser Met Asp Phe Ala Ile Ser Lys Leu Lys Lys Ala
114                405                410                415
116 Gly Asn Gln Thr Gly Leu Tyr Val Leu Arg Cys Ser Pro Lys Asp Phe
117                420                425                430
119 Asn Lys Tyr Phe Leu Thr Phe Ala Val Glu Arg Glu Asn Val Ile Glu
120                435                440                445
122 Tyr Lys His Cys Leu Ile Thr Lys Asn Glu Asn Glu Glu Tyr Asn Leu
123                450                455                460
125 Ser Gly Thr Lys Lys Asn Phe Ser Ser Leu Lys Asp Leu Leu Asn Cys
126 465                470                475                480
128 Tyr Gln Met Glu Thr Val Arg Ser Asp Asn Ile Ile Phe Gln Phe Thr
129                485                490                495
131 Lys Cys Cys Pro Pro Lys Pro Lys Asp Lys Ser Asn Leu Leu Val Phe
132                500                505                510
134 Arg Thr Asn Gly Val Ser Asp Val Pro Thr Ser Pro Thr Leu Gln Arg
135                515                520                525
137 Pro Thr His Met Asn Gln Met Val Phe His Lys Ile Arg Asn Glu Asp

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/580,458A

DATE: 08/22/2006

TIME: 10:54:30

Input Set : A:\65691-445CorrSeqList.txt

Output Set: N:\CRF4\08222006\J580458A.raw

```

138      530      535      540
140 Leu Ile Phe Asn Glu Ser Leu Gly Gln Gly Thr Phe Thr Lys Ile Phe
141 545      550      555      560
143 Lys Gly Val Arg Arg Glu Val Gly Asp Tyr Gly Gln Leu His Glu Thr
144      565      570      575
146 Glu Val Leu Leu Lys Val Leu Asp Lys Ala His Arg Asn Tyr Ser Glu
147      580      585      590
149 Ser Phe Phe Glu Ala Ala Ser Met Met Ser Lys Leu Ser His Lys His
150      595      600      605
152 Leu Val Leu Asn Tyr Gly Val Cys Phe Cys Gly Asp Glu Asn Ile Leu
153 610      615      620
155 Val Gln Glu Phe Val Lys Phe Gly Ser Leu Asp Thr Tyr Leu Lys Lys
156 625      630      635      640
158 Asn Lys Asn Cys Ile Asn Ile Leu Trp Lys Leu Glu Val Ala Lys Gln
159      645      650      655
161 Leu Ala Trp Ala Met His Phe Leu Glu Glu Asn Thr Leu Ile His Gly
162      660      665      670
164 Asn Val Cys Ala Lys Asn Ile Leu Leu Ile Arg Glu Glu Asp Arg Lys
165      675      680      685
167 Thr Gly Asn Pro Pro Phe Ile Lys Leu Ser Asp Pro Gly Ile Ser Ile
168      690      695      700
170 Thr Val Leu Pro Lys Asp Ile Leu Gln Glu Arg Ile Pro Trp Val Pro
171 705      710      715      720
173 Pro Glu Cys Ile Glu Asn Pro Lys Asn Leu Asn Leu Ala Thr Asp Lys
174      725      730      735
176 Trp Ser Phe Gly Thr Thr Leu Trp Glu Ile Cys Ser Gly Gly Asp Lys
177      740      745      750
179 Pro Leu Ser Ala Leu Asp Ser Gln Arg Lys Leu Gln Phe Tyr Glu Asp
180      755      760      765
182 Arg His Gln Leu Pro Ala Pro Lys Trp Ala Glu Leu Ala Asn Leu Ile
183      770      775      780
185 Asn Asn Cys Met Asp Tyr Glu Pro Asp Phe Arg Pro Ser Phe Arg Ala
186 785      790      795      800
188 Ile Ile Arg Asp Leu Asn Ser Leu Phe Thr Pro Asp Tyr Glu Leu Leu
189      805      810      815
191 Thr Glu Asn Asp Met Leu Pro Asn Met Arg Ile Gly Ala Leu Gly Phe
192      820      825      830
194 Ser Gly Ala Phe Glu Asp Arg Asp Pro Thr Gln Phe Glu Glu Arg His
195      835      840      845
197 Leu Lys Phe Leu Gln Gln Leu Gly Lys Gly Asn Phe Gly Ser Val Glu
198      850      855      860
200 Met Cys Arg Tyr Asp Pro Leu Gln Asp Asn Thr Gly Glu Val Val Ala
201 865      870      875      880
203 Val Lys Lys Leu Gln His Ser Thr Glu Glu His Leu Arg Asp Phe Glu
204      885      890      895
206 Arg Glu Ile Glu Ile Leu Lys Ser Leu Gln His Asp Asn Ile Val Lys
207      900      905      910
209 Tyr Lys Gly Val Cys Tyr Ser Ala Gly Arg Arg Asn Leu Lys Leu Ile
210      915      920      925

```

RAW SEQUENCE LISTING

DATE: 08/22/2006

PATENT APPLICATION: US/10/580,458A

TIME: 10:54:30

Input Set : A:\65691-445CorrSeqList.txt

Output Set: N:\CRF4\08222006\J580458A.raw

```

212 Met Glu Tyr Leu Pro Tyr Gly Ser Leu Arg Asp Tyr Leu Gln Lys His
213      930      935      940
215 Lys Glu Arg Ile Asp His Ile Lys Leu Leu Gln Tyr Thr Ser Gln Ile
216 945      950      955      960
218 Cys Lys Gly Met Glu Tyr Leu Gly Thr Lys Arg Tyr Ile His Arg Asp
219      965      970      975
221 Leu Ala Thr Arg Asn Ile Leu Val Glu Asn Glu Asn Arg Val Lys Ile
222      980      985      990
224 Gly Asp Phe Gly Leu Thr Lys Val Leu Pro Gln Asp Lys Glu Tyr Tyr
225      995      1000      1005
227 Lys Val Lys Glu Pro Gly Glu Ser Pro Ile Phe Trp Tyr Ala Pro
228      1010      1015      1020
230 Glu Ser Leu Thr Glu Ser Lys Phe Ser Val Ala Ser Asp Val Trp
231      1025      1030      1035
233 Ser Phe Gly Val Val Leu Tyr Glu Leu Phe Thr Tyr Ile Glu Lys
234      1040      1045      1050
236 Ser Lys Ser Pro Pro Ala Glu Phe Met Arg Met Ile Gly Asn Asp
237      1055      1060      1065
239 Lys Gln Gly Gln Met Ile Val Phe His Leu Ile Glu Leu Leu Lys
240      1070      1075      1080
242 Asn Asn Gly Arg Leu Pro Arg Pro Asp Gly Cys Pro Asp Glu Ile
243      1085      1090      1095
245 Tyr Met Ile Met Thr Glu Cys Trp Asn Asn Asn Val Asn Gln Arg
246      1100      1105      1110
248 Pro Ser Phe Arg Asp Leu Ala Leu Arg Val Asp Gln Ile Arg Asp
249      1115      1120      1125
251 Asn Met Ala Gly
252      1130
254 <210> SEQ ID NO: 2
255 <211> LENGTH: 5097
256 <212> TYPE: DNA
257 <213> ORGANISM: homo sapiens
260 <220> FEATURE:
261 <223> OTHER INFORMATION: G1849T mutation in jak2 gene
263 <400> SEQUENCE: 2
264 ctgcaggaag gagagaggaa gaggagcaga agggggcagc agcggacgcc gctaacggcc      60
266 tccctcggcg ctgacaggct gggccggcgc ccggctcgct tgggtgttcg cgtcgccact      120
268 tcggcttctc ggccggtcgg gccctcggc ccgggcttgc ggcgcgcgtc ggggctgagg      180
270 gctgctgcgg cgcagggaga ggcctggtcc tcgctgccga gggatgtgag tgggagctga      240
272 gccacactg gagggccccc gagggcccag cctggaggtc gttcagagcc gtgcccggcc      300
274 cggggcttcg cagaccttga cccgccgggt aggagccgcc cctgcgggct cgagggcgcg      360
276 ctctggtcgc ccgatctgtg tagccggttt cagaagcagg caacaggaac aagatgtgaa      420
278 ctgtttctct tctgcagaaa aagaggctct tcctcctcct cccgcgacgg caaatgttct      480
280 gaaaaagact ctgcatggga atggcctgcc ttacgatgac agaaatggag ggaacatcca      540
282 cctcttctat atatcagaat ggtgatattt ctggaaatgc caattctatg aagcaaatag      600
284 atccagttct tcaggtgtat ctttaccatt cccttgggaa atctgaggca gattatctga      660
286 cctttccatc tggggagtat gttgcagaag aaatctgtat tgctgcttct aaagcttgtg      720
288 gtatcacacc tgtgtatcat aatatgtttg ctttaatgag tgaaacagaa aggatctggt      780
290 atccacccaa ccatgtcttc catatagatg agtcaaccag gcataatgta ctctacagaa      840

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/580,458A

DATE: 08/22/2006

TIME: 10:54:30

Input Set : A:\65691-445CorrSeqList.txt

Output Set: N:\CRF4\08222006\J580458A.raw

292	taagatttta	ctttcctcgt	tggatttgca	gtggcagcaa	cagagcctat	cggcatggaa	900
294	tatctcgagg	tgctgaagct	cctcttcttg	atgactttgt	catgtcttac	ctctttgctc	960
296	agtggcggca	tgattttgtg	cacggatgga	taaaagtacc	tgtgactcat	gaaacacagg	1020
298	aagaatgtct	tgggatggca	gtgttagata	tgatgagaat	agccaaagaa	aacgatcaaa	1080
300	ccccactggc	catctataac	tctatcagct	acaagacatt	cttaccaaaa	tgtattcgag	1140
302	caaagatcca	agactatcat	attttgacaa	ggaagcgaat	aaggtacaga	tttcgcagat	1200
304	ttattcagca	attcagccaa	tgcaaagcca	ctgccagaaa	cttgaaactt	aagtatctta	1260
306	taaatctgga	aactctgcag	tctgccttct	acacagagaa	atltgaagta	aaagaacctg	1320
308	gaagtgggtc	ttcaggtgag	gagatttttg	caaccattat	aataactgga	aacggtggaa	1380
310	ttcagtggtc	aagagggaaa	cataaagaaa	gtgagacact	gacagaacag	gatttacagt	1440
312	tatattgcca	ttttccta	attattgatg	tcagtattaa	gcaagcaaac	caagagggtt	1500
314	caaataagaa	ccgagttgta	actatccata	agcaagatgg	taaaaatctg	gaaattgaac	1560
316	ttagctcatt	aagggaagct	ttgtctttcg	tgtcattaat	tgatggatat	tatagattaa	1620
318	ctgcagatgc	acatcattac	ctctgtaaa	aagtagcacc	tccagccgtg	cttgaaaata	1680
320	tacaaagcaa	ctgtcatggc	ccaatttcga	tggattttgc	cattagtaaa	ctgaagaaag	1740
322	caggtaatca	gactggactg	tatgtacttc	gatgcagtc	taaggacttt	aataaatatt	1800
324	ttttgacttt	tgtctgcgag	cgagaaaatg	tcattgaata	taaacactgt	ttgattacaa	1860
326	aaaatgagaa	tgaagagtac	aacctcagtg	ggacaaagaa	gaacttcagc	agtcttaaa	1920
328	atcttttgaa	ttgttaccag	atggaaactg	ttcgtcaga	caatataatt	ttccagttta	1980
330	ctaaatgctg	tcccccaaag	ccaaaagata	aatcaaacct	tctagtcttc	agaacgaatg	2040
332	gtgtttctga	tgtaccaacc	tcaccaacat	tacagaggcc	tactcatatg	aaccaaatgg	2100
334	tgtttcacaa	aatcagaaat	gaagatttga	tatttaata	aagccttggc	caaggcactt	2160
336	ttacaaagat	ttttaaaggc	gtacgaagag	aagtaggaga	ctacgggtcaa	ctgcatgaaa	2220
338	cagaagttct	tttaaaagtt	ctggataaag	cacacagaaa	ctattcagag	tctttctttg	2280
340	aagcagcaag	tatgatgagc	aagctttctc	acaagcattt	ggtttttaaat	tatggagtat	2340
342	gtttctgtgg	agacgagaat	attctgggtc	aggagtttgt	aaaatttgga	tcactagata	2400
344	catatctgaa	aaagaataaa	aattgtataa	atatattatg	gaaacttgaa	gttgctaaac	2460
346	agttggcatg	ggccatgcat	tttctagaag	aaaacaccct	tattcatggg	aatgtatgtg	2520
348	ccaaaaatat	tctgcttatt	agagaagaag	acaggaagac	aggaaatcct	cctttcatca	2580
350	aacttagtga	tcctggcatt	agtattacag	ttttgccaaa	ggacattctt	caggagagaa	2640
352	taccatgggt	accacctgaa	tgcattgaaa	atcctaataa	tttaaatgtg	gcaacagaca	2700
354	aatggagttt	tgttaccact	ttgtgggaaa	tctgcagtgg	aggagataaa	cctctaagtg	2760
356	ctctggattc	tcaaaagaa	ctacaatttt	atgaagatag	gcatcagctt	cctgcaccaa	2820
358	agtgggcaga	attagcaaac	cttataaata	attgtatgga	ttatgaacca	gatttcaggc	2880
360	cttctttcag	agccatcata	cgagatctta	acagtttgtt	tactccagat	tatgaactat	2940
362	taacagaaaa	tgacatgtta	ccaaatatga	ggatagggtg	cctagggttt	tctgggtgct	3000
364	ttgaagaccg	ggatcctaca	cagtttgaag	agagacattt	gaaatttcta	cagcaacttg	3060
366	gcaagggtaa	ttttgggagt	gtggagatgt	gccggtatga	ccctctacag	gacaacactg	3120
368	gggaggtgg	cgctgtaaaa	aagcttcagc	atagtactga	agagcaccta	agagactttg	3180
370	aaagggaaat	tgaaatcctg	aaatccctac	agcatgacaa	cattgtaaa	tacaagggag	3240
372	tgtgctacag	tgctggctcg	cgtaatctaa	aattaattat	ggaatattta	ccatatggaa	3300
374	gtttacgaga	ctatcttcaa	aaacataaag	aacggataga	tcacataaaa	cttctgcagt	3360
376	acacatctca	gatatgcaag	ggtatggagt	atcttggtac	aaaaaggtat	atccacaggg	3420
378	atctggcaac	gagaaatata	ttggtggaga	acgagaacag	agttaaaatt	ggagattttg	3480
380	ggttaaccaa	agtcttgcca	caagacaaag	aataactata	agtaaaagaa	cctgggtgaa	3540
382	gtcccatatt	ctggatgct	ccagaatcac	tgacagagag	caagttttct	gtggcctcag	3600
384	atgtttggag	ctttggagtg	gttctgtatg	aacttttcac	atacattgag	aagagtaaaa	3660
386	gtccaccagc	ggaatttatg	cgtatgattg	gcaatgacaa	acaaggacag	atgatcgtgt	3720
388	tccatttgat	agaacttttg	aagaataatg	gaagattacc	aagaccagat	ggatgcccg	3780

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/580,458A

DATE: 08/22/2006
TIME: 10:54:31

Input Set : A:\65691-445CorrSeqList.txt
Output Set: N:\CRF4\08222006\J580458A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:13; N Pos. 20,21

Seq#:14; N Pos. 20,21

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/580,458A

DATE: 08/22/2006

TIME: 10:54:31

Input Set : A:\65691-445CorrSeqList.txt

Output Set: N:\CRF4\08222006\J580458A.raw

L:20 M:270 C: Current Application Number differs, Replaced Current Application Number

L:21 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:599 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:0

L:622 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:0